

MinION-based sequencing of pathogens



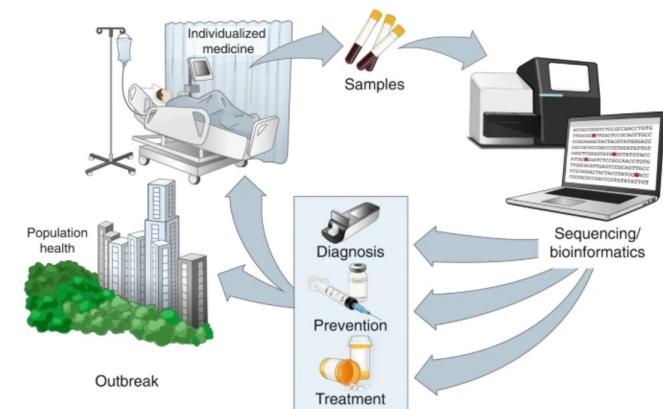
Susana.campino@lshtm.ac.uk

Sequencing of pathogen genomes

Technological advances are enabling the application of pathogen genome sequencing. Whole-genome sequencing of many pathogens can now be done directly from clinical samples and in (almost) real time during an outbreak.

Genomic studies of pathogens helps to understand:

- biology
- transmission
- pathogenesis
- host-pathogen interactions
- evolution (e.g. signatures of drug resistance)
- epidemiology



Genomic studies have been applied to:

- development of rapid diagnostic tools
- surveillance of pathogens, relatedness, drug resistance
- identification of sources of infection and transmission trees
- development of vaccines and drugs

Overview on Whole genome sequencing technologies

Second (Next) Generation sequencing technology

Multiple copies of small DNA fragments are sequenced



Illumina systems
High-low throughput systems



Ion Torrent PGM system
Medium/low throughput system



Roche/454 sequencing
Medium/low throughput system



PacBio sequencing
Low throughput system



Oxford Nanopore (MinION)
Low/Medium throughput system



SmidgION

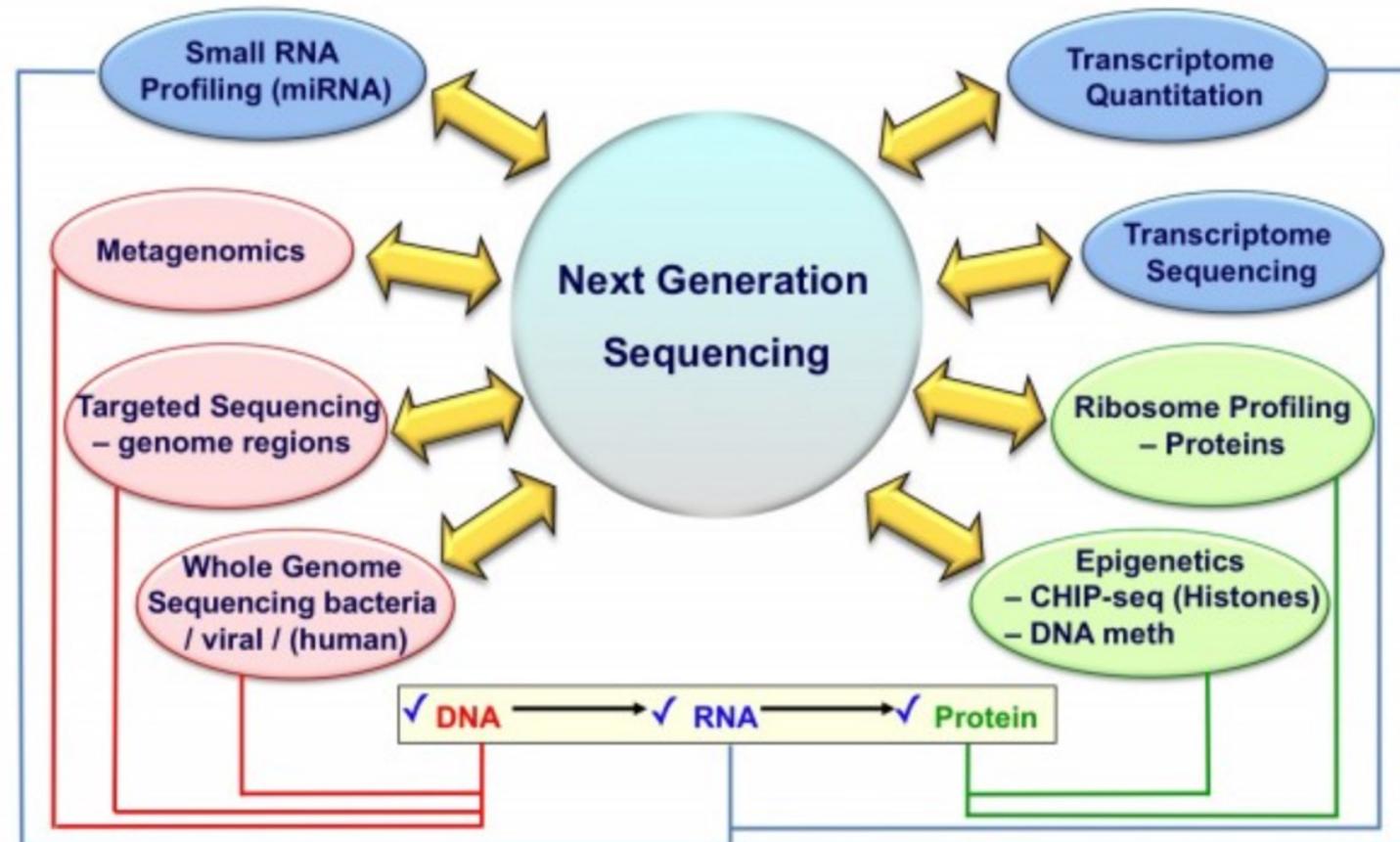
Different Nanopore platforms



Mk1C with fully integrated computer



Next-generation sequencing applications



Whole genome sequencing (WGS) approach

Although differing in their chemistries and processes, the platforms have broadly similar workflows.

Second Generation

- 1- Total DNA is cut into small fragments (~300bp) randomly
- 2-Adapters added to each fragment
- 3- Massive parallel amplification of fragments
- 4- Sequencing of clonal fragments



Analyses involves assembling the fragments to reconstruct the original DNA (like a puzzle).

Third Generation

- 1- DNA kept as long fragments
- 2-Adapters added to each fragment
- 3- Real time sequencing of each fragments



Important factors to select the type of platform include:

- How many samples/ cost per run
- sample preparation complexity
- run time
- simplicity of data analysis and read lengths generated.

Whole genome sequencing (WGS) approach

Although differing in their chemistries and processes, the platforms have broadly similar workflows.

Second Generation

- 1- Total DNA is cut into small fragments (~300bp) randomly
- 2-Adapters added to each fragment
- 3- Massive parallel amplification of fragments
- 4- Sequencing of clonal fragments



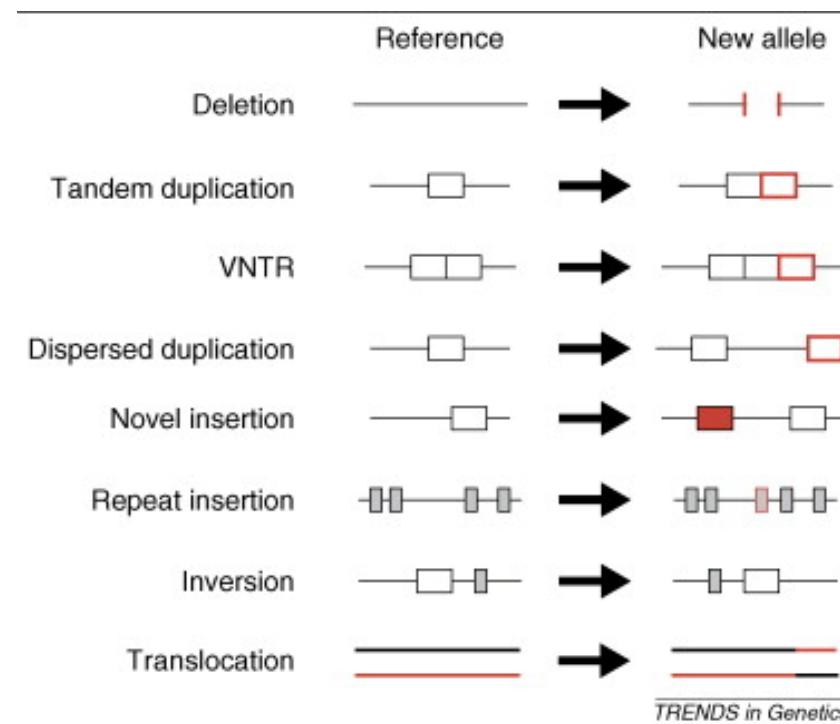
If the picture puzzle is not known

Analyses involves
De novo assembly

Advantage of long read sequencing

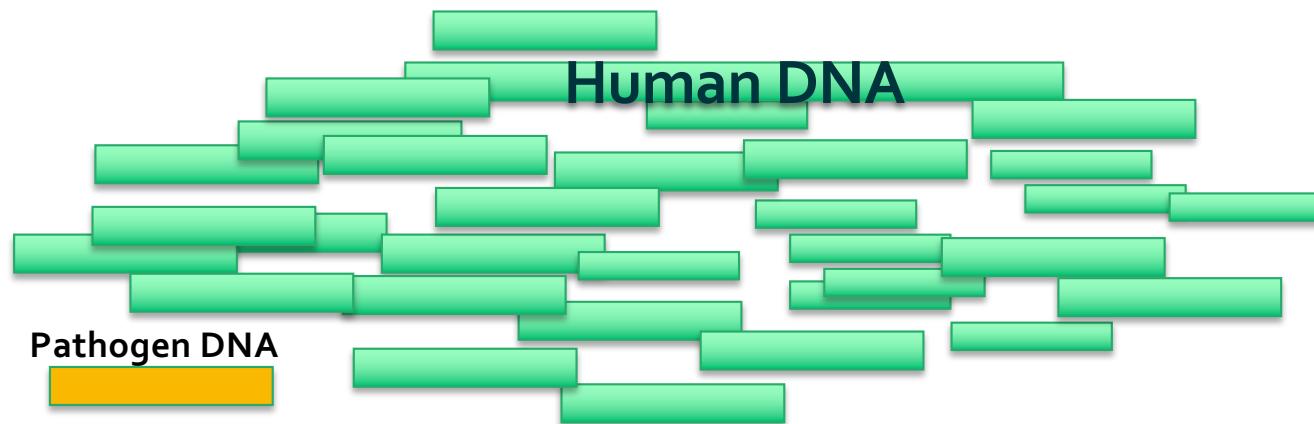
- Easier to assemble genomes
- improve de novo assembly (telomere – to- telomere)
- Obtain long transcript, helping isoform identification
- Detection of structural variants.

Structural variants (SVs) are estimated to contribute to a great proportion of genomic differences between individuals:



WGS: Problem with clinical samples

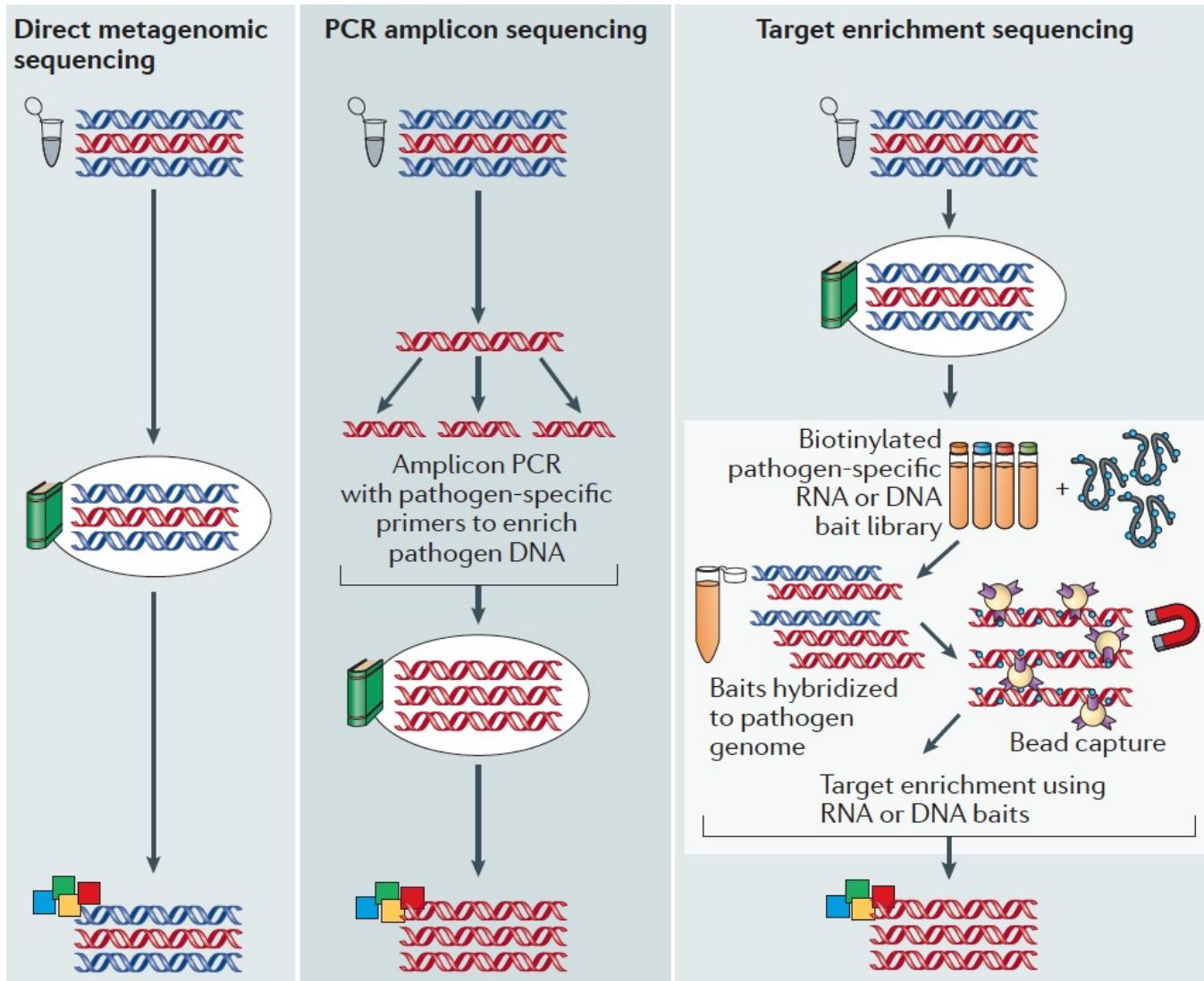
Clinical sample= human DNA + pathogen DNA
 $>99.9\%$ $< 0.1\%$



Genomic data will have mainly human fragments and will be difficult to analyze the pathogen genome

Approaches to sequence the pathogen genomes from clinical samples

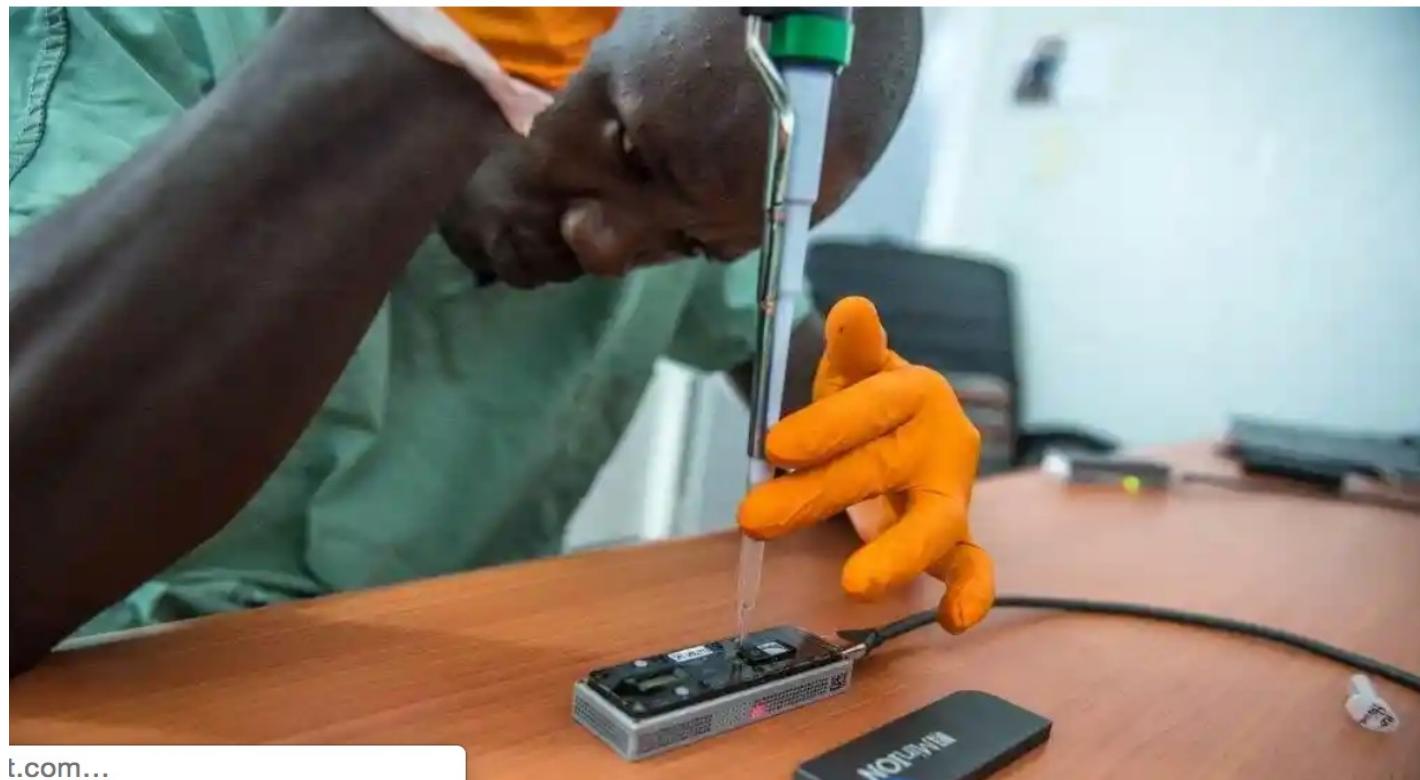
Clinical samples have **host DNA contamination** and it is therefore difficult to sequence the pathogen genome



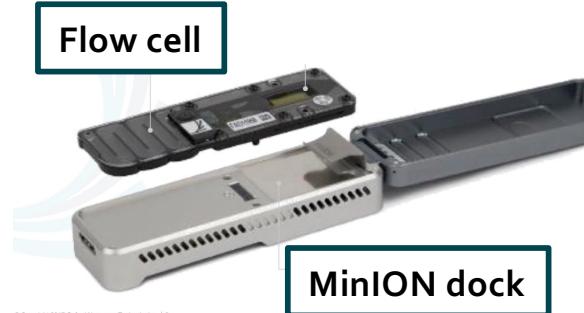
Organize a mobile lab to support outbreaks

From Ebola to Zika, tiny mobile lab gives real-time DNA data on outbreaks

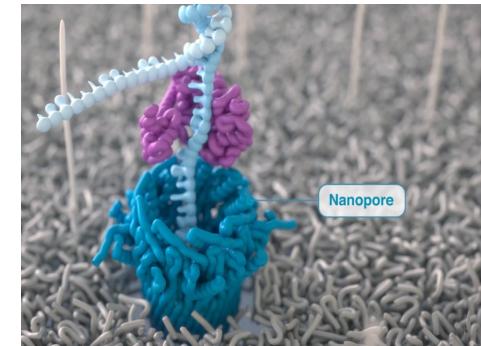
A genomic surveillance system which fits in a suitcase can help health workers to quickly understand the spread of viruses and break the chain of infection



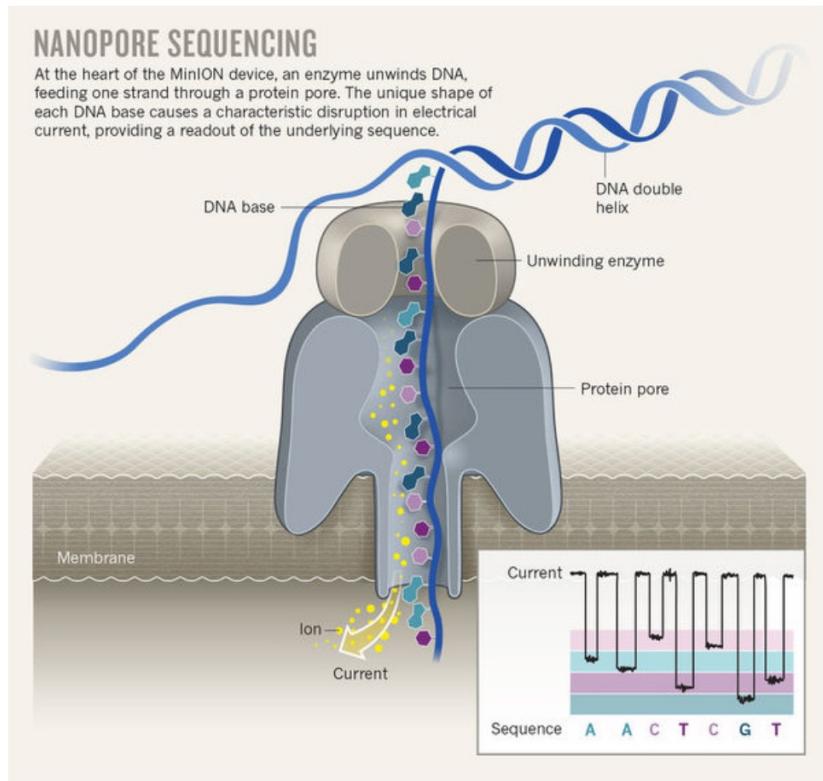
Overview: Nanopore sequencing technology



In the sensor array there are ~2000 wells each containing a single nanopore



Overview: Nanopore sequencing technology



Broad Explanation of pore base-calling method

An enzyme unzips the DNA and one strand passes into the pore

A flow of ions passes in the pore

When the DNA passes through the pore it affects the ionic current

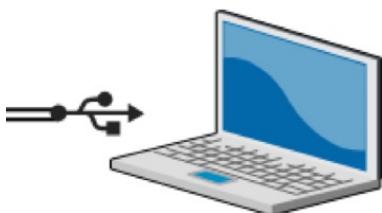
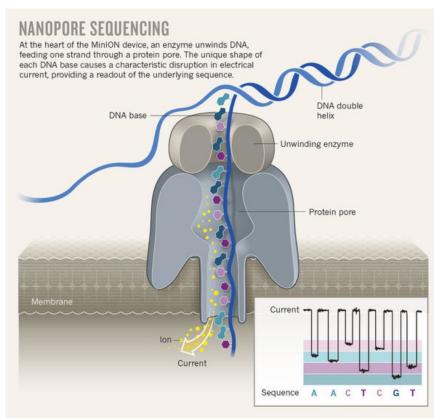
The current is changed as the bases G, A, T and C pass through the pore in different combinations.

Changes segmented as discrete events that have an associated duration, mean amplitude, and variance.

These signals can be used to infer computationally the DNA sequence.

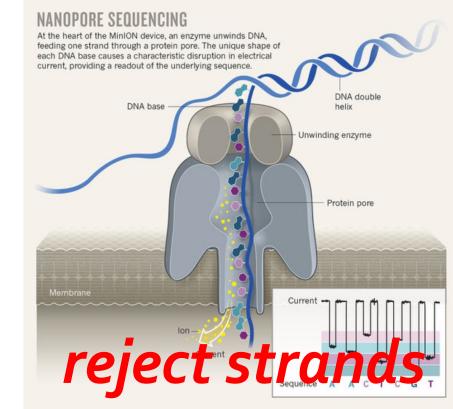
Adaptive sequencing

A type of selective sequencing

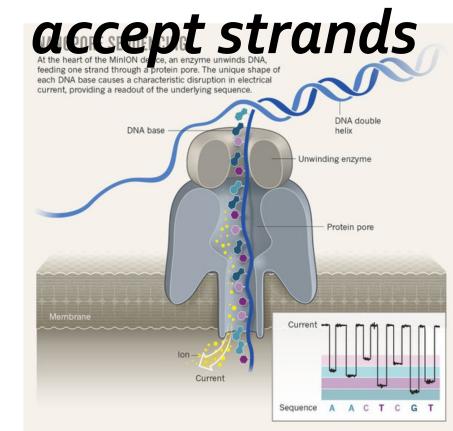


The user can programme their system to accept or reject strands based on a configuration specified in software.

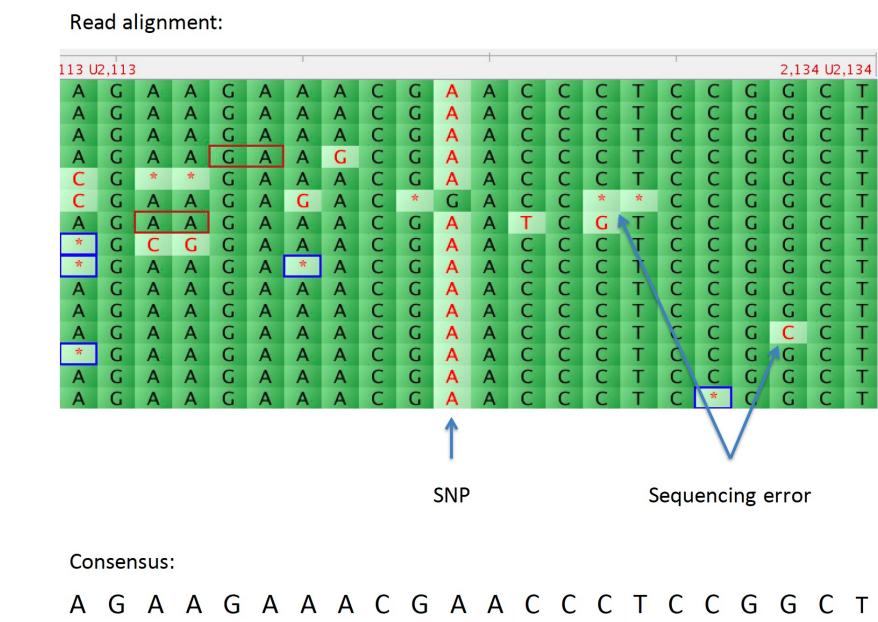
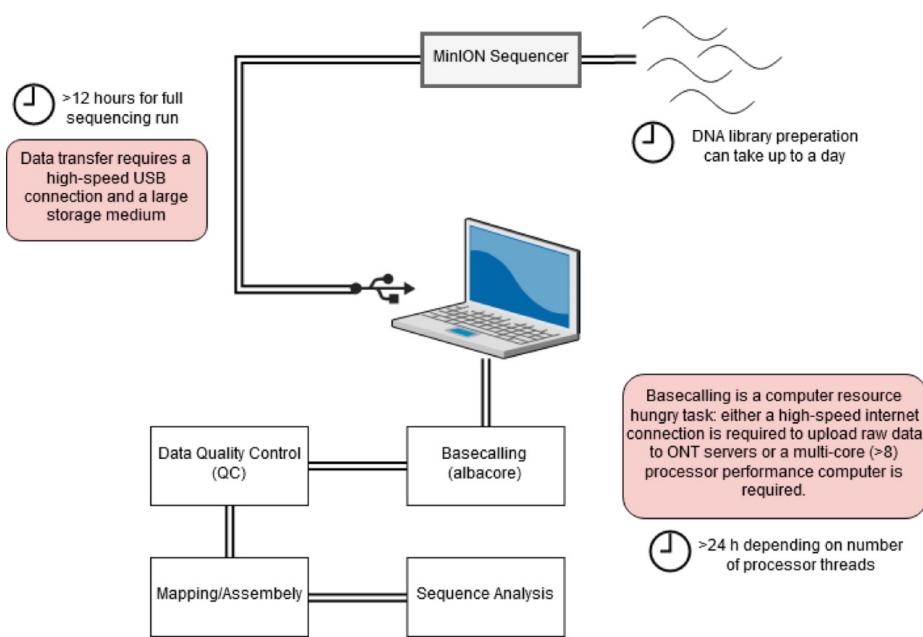
Human DNA in pore



Pathogen DNA in pore



Sequencing pipeline



MinION has a higher error rates in the base called reads

(10% as compared to 1% for Illumina)

Organising a mobile lab with MinION

In remote areas logistical and technical challenges can be significant.

Nanopore sequencing can easily and quickly be deployed into remote areas, and allows sequencing to be done directly in an outbreak area.

General lab material

Gloves, lab coat, pipettes, tips, nuclease free water, Ethanol

Equipment and material for DNA/ RNA extraction

Kits for rapid DNA extraction

Eppendorf 1.5ml tubes

Minicentrifuge

Mini heatblock

Equipment and material for target amplicon of pathogens

PCR tubes and caps

Reagents for isothermal PCR (one step PCR)

Equipment and material for library preparation and sequencing

Kits for sequencing

Magnetic beads and rack

MinION flow cell

MinION

Laptop

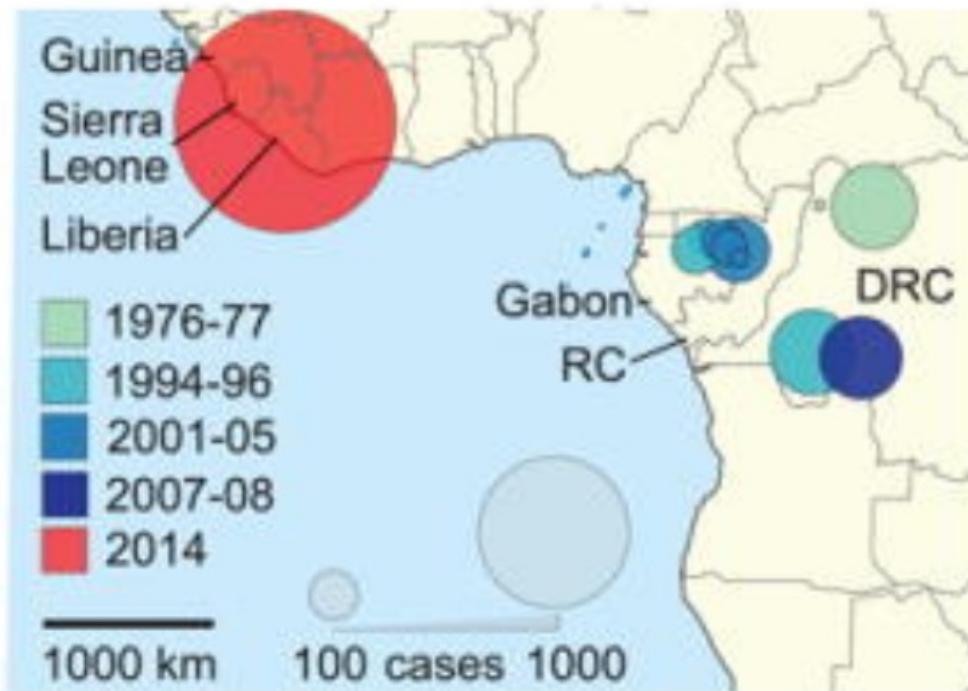


Sequencing of Ebola Virus Genomes Using Nanopore Technology



Ebola virus is a lethal human pathogen, causing Ebola virus disease with an average case fatality rate of 78%.

Previous outbreaks were confined to remote regions of central Africa; the largest, in 1976, had 318 cases. The outbreak of 2014 started in February 2014 in Guinea, West Africa and spread into Liberia in March, Sierra Leone in May, and Nigeria in late July.



Science. 2014 Sep 12;345(6202):1369-72.

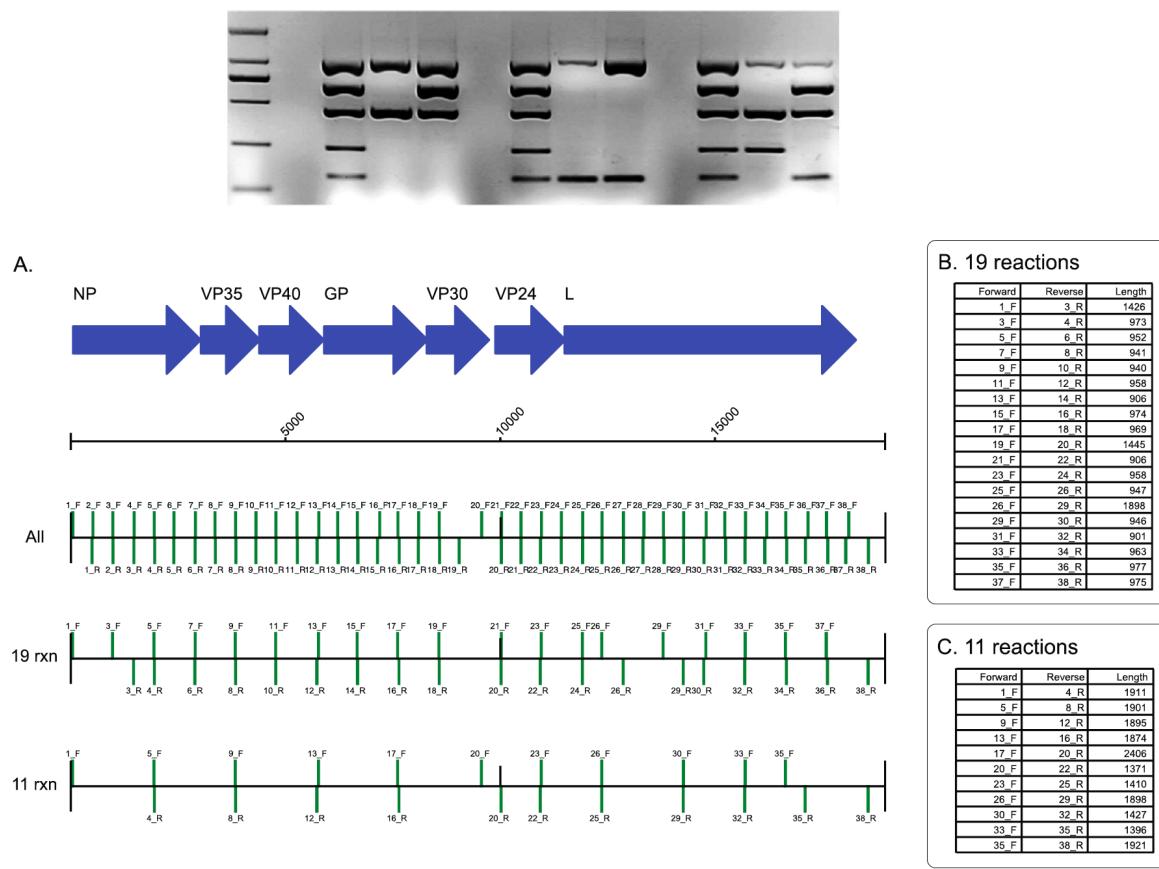
Sequencing of Ebola Virus Genomes Using Nanopore Technology

Deployment of the portable genome surveillance system in Guinea



Sequencing of Ebola Virus Genomes Using Nanopore Technology

PCR amplicon sequencing of whole Ebola virus genome



Sequencing of Ebola Virus Genomes Using Nanopore Technology



Automated design of PCR amplicon sequencing of whole genome

The screenshot shows the homepage of the primalscheme website. The header features the text "primalscheme" in large white letters and "primer panels for multiplex PCR" in smaller yellow text. A large blue 3D dodecahedron icon with the number "00" on it is positioned on the right side of the header. Below the header, the main form area has a red background. It contains fields for "Design a new scheme" (with a "Reset defaults" button), "Amplicon size" (set to 400), and "Scheme name" (e.g., nCov-400). On the left, there are "Options" for "High-GC mode" and "Pinned", both of which are currently off. A note below these options says "Use the standard protocol for these settings." At the bottom, there is a "Design scheme" button with a play icon and a "View demo inputs" link.

primalscheme
primer panels for multiplex PCR

Design a new scheme Reset defaults

Options

High-GC mode

Pinned

Use the standard protocol for these settings.

Amplicon size
400

Min/max will be set at 5% either side of target.

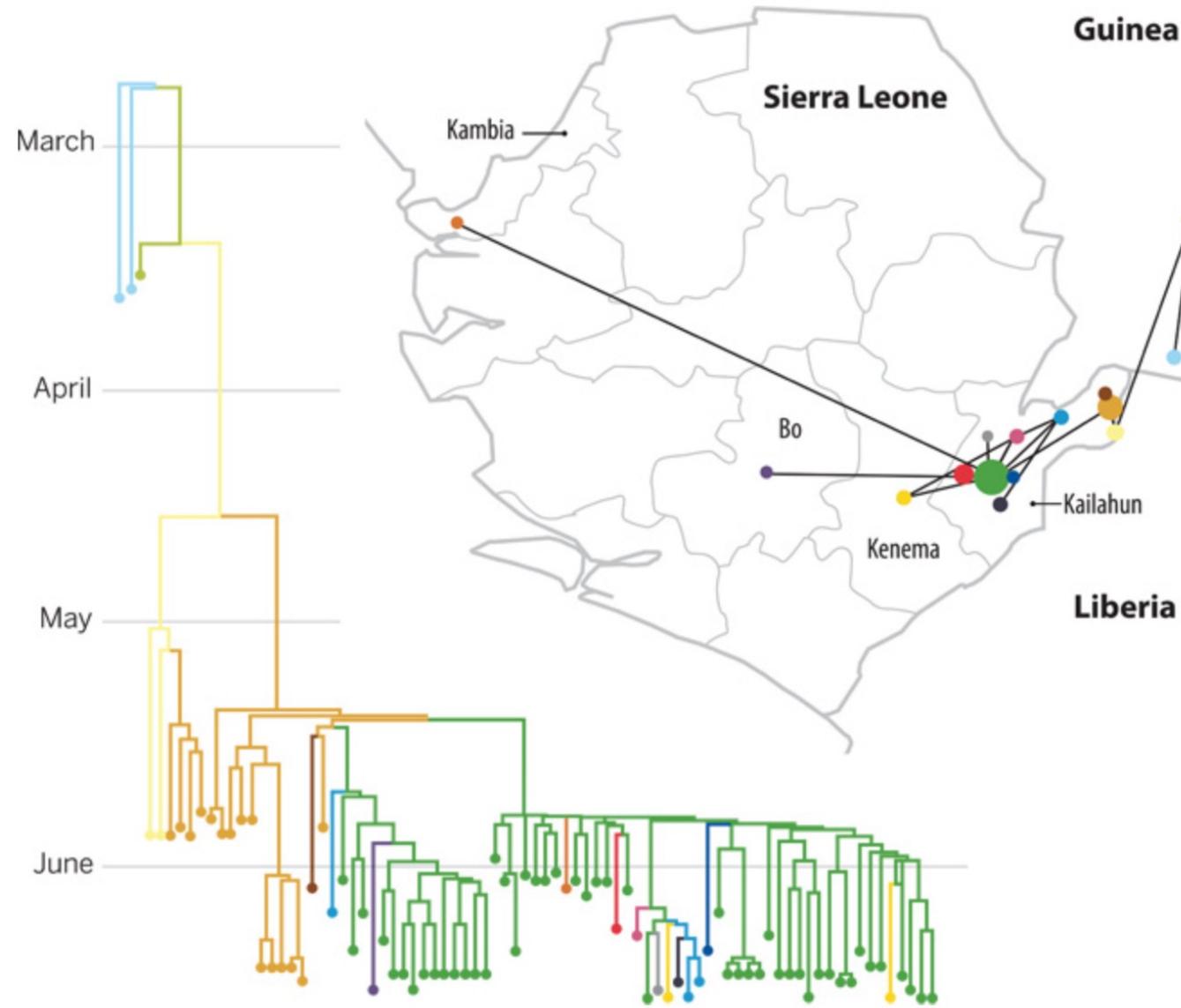
Scheme name
e.g., nCov-400

A short name/prefix for your scheme, no spaces.

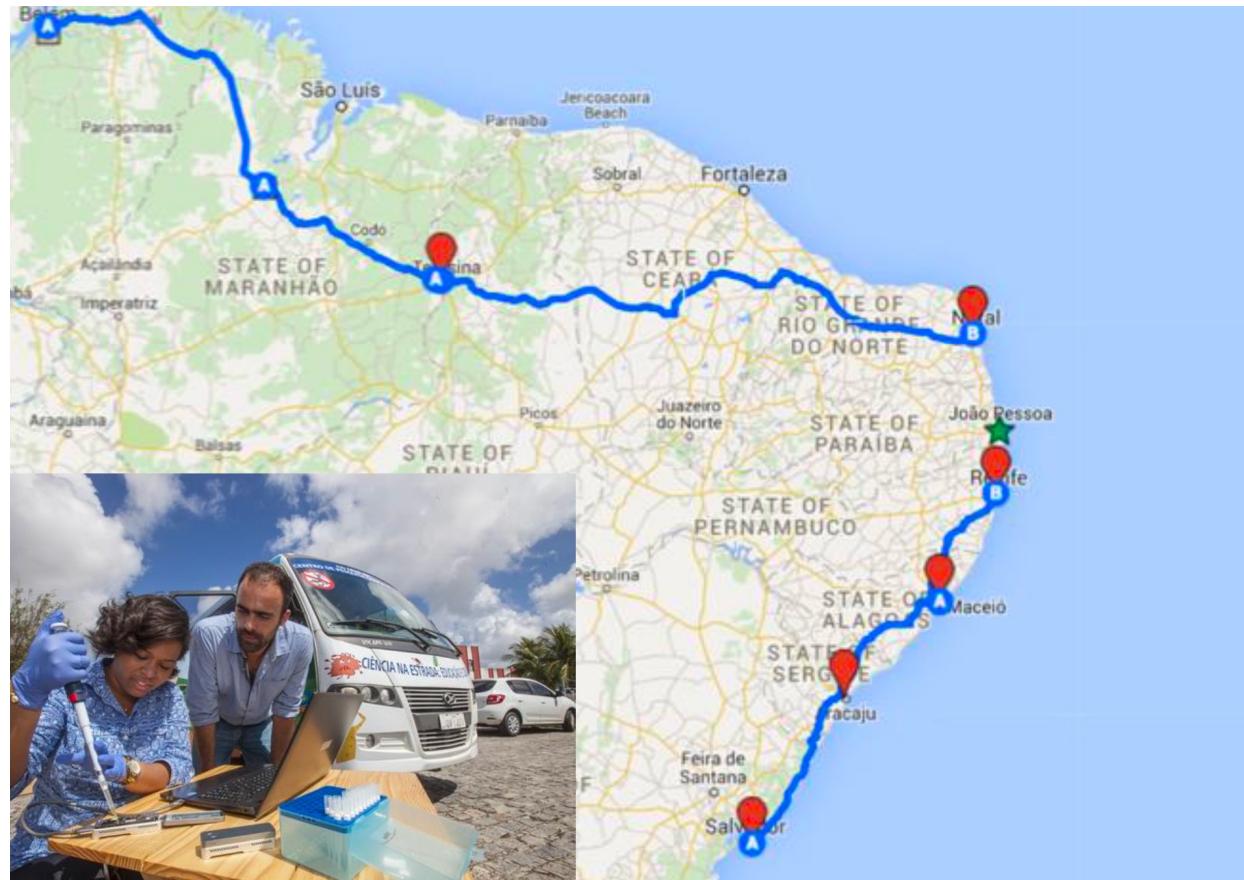
[View demo inputs](#)

<https://primalscheme.com/>

Genomics informs Ebola outbreak

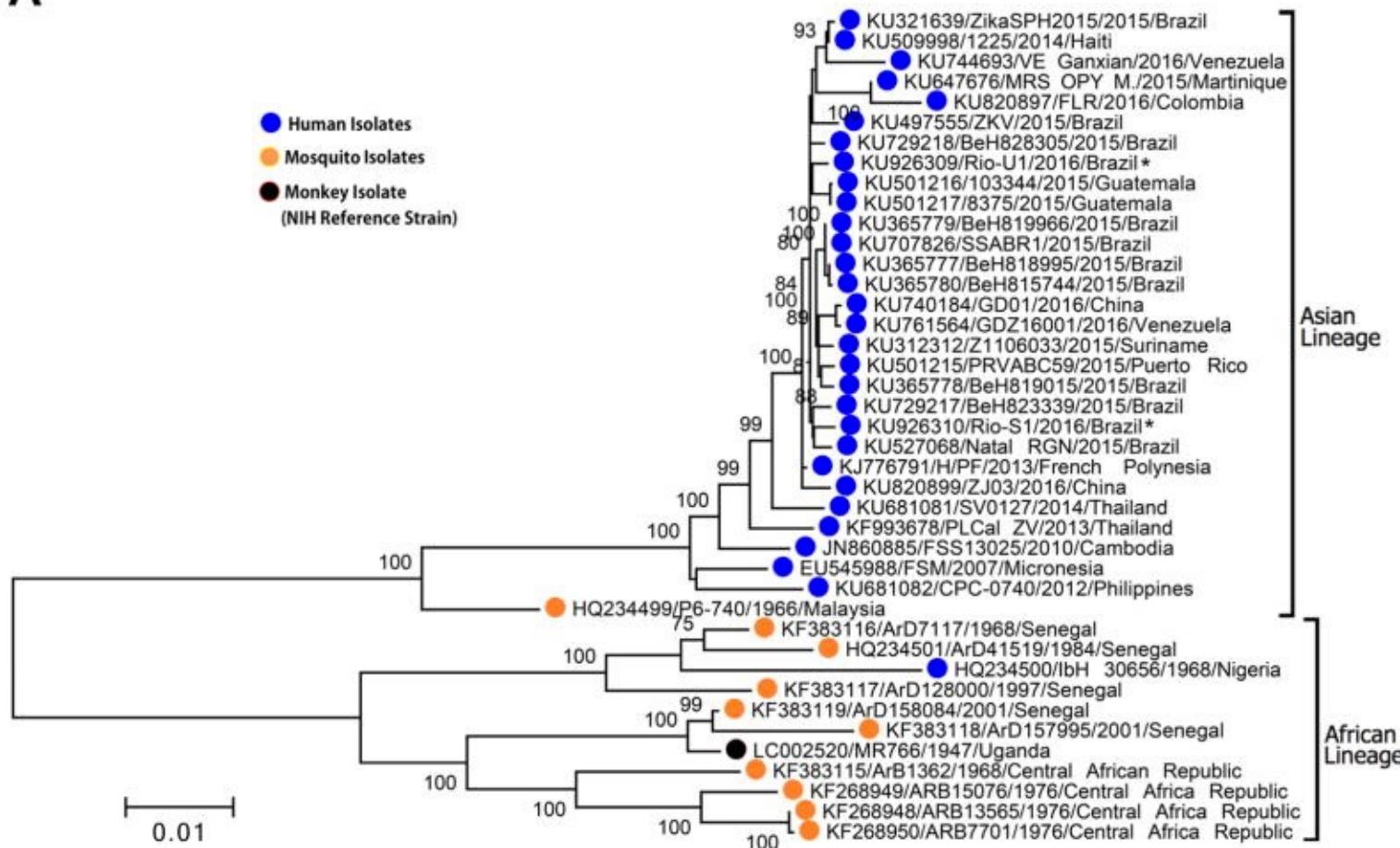


Biolabs bring genomics to the field



Virus genomics informs Zika outbreak in Brazil

A



Oxford Nanopore used during Pandemic

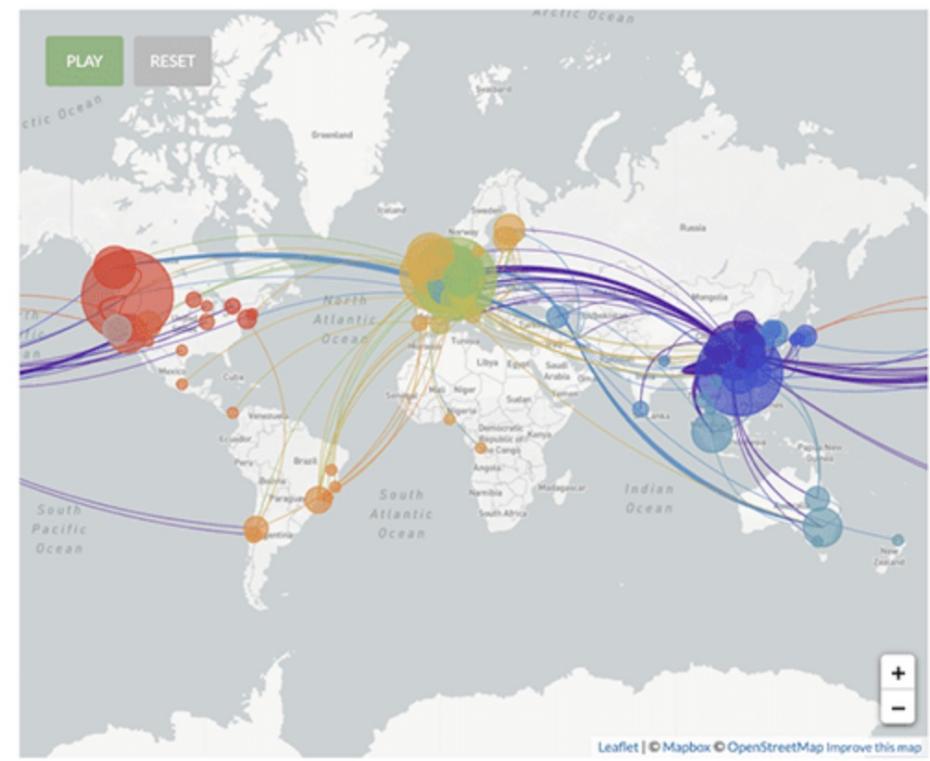
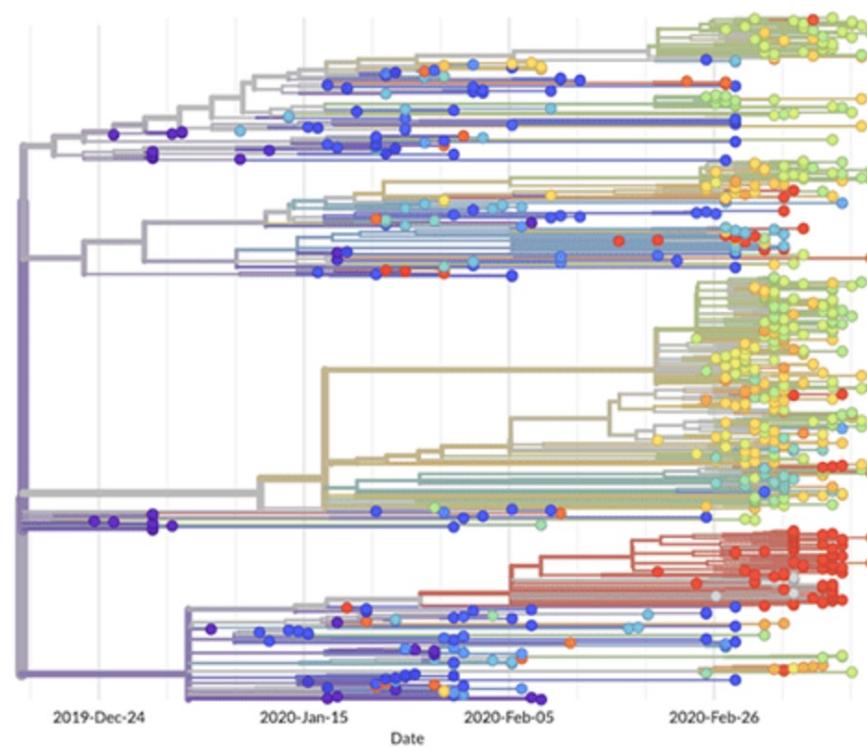


Phylo-genomic-temporal infection tracking of SARS CoV-2

Genomics has been an important tool to support SARs-Cov2 outbreak

- Virus identification
- Transmission chains
- Global Spread
- Diagnostic tools
- Vaccine development

Real-time tracking of pathogen evolution



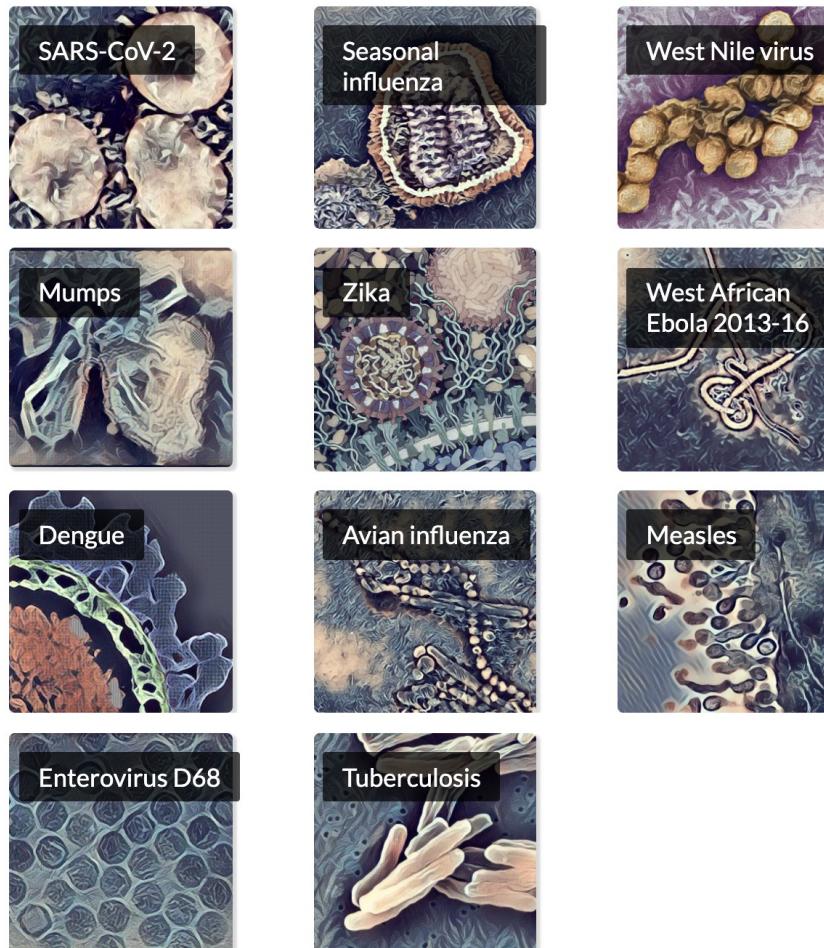
Nextstrain, enabled by data from GISAID

Phylo-genomic-temporal infection tracking

Nextstrain Real-time tracking of pathogen evolution

Explore pathogens

Genomic analyses of specific pathogens kept up-to-date by the Nextstrain team



Nextstrain, enabled by data from GISAID